

Original Paper

Comparing In-Person, Standard Telehealth, and Remote Musculoskeletal Examination With a Novel Augmented Reality Exercise Game System: Pilot Cross-Sectional Comparison Study

Richard Wu^{1,2}, MD; Keerthana Chakka^{1,3}, MD; Sara Belko⁴, MS; Ninad Khargonkar⁵, MS; Kevin Desai⁶, PhD; Balakrishnan Prabhakaran⁵, PhD; Thiru Annaswamy⁷, MD, MA

¹The University of Texas Southwestern Medical Center, Dallas, TX, United States

²Johns Hopkins Hospital, Baltimore, MD, United States

³Northwestern Medicine, Chicago, IL, United States

⁴Sidney Kimmel Medical College at Thomas Jefferson University, Philadelphia, PA, United States

⁵Department of Computer Science, The University of Texas at Dallas, Richardson, TX, United States

⁶Department of Computer Science, The University of Texas at San Antonio, San Antonio, TX, United States

⁷Penn State Health Milton S. Hershey Medical Center, Penn State University College of Medicine, Hershey, PA, United States

Corresponding Author:

Thiru Annaswamy, MD, MA
Penn State Health Milton S. Hershey Medical Center
Penn State University College of Medicine
500 University Drive, PO Box 850, Mail Code A101
Hershey, PA, 17033
United States
Phone: 717-531-4263
Email: tannaswamy@pennstatehealth.psu.edu

Abstract

Background: Current telemedicine technologies are not fully optimized for conducting physical examinations. The Virtual Remote Tele-Physical Examination (VIRTEPEX) system, a novel proprietary technology platform using a Microsoft Kinect-based augmented reality game system to track motion and estimate force, has the potential to assist with conducting asynchronous, remote musculoskeletal examinations.

Objective: This pilot study evaluated the feasibility of the VIRTEPEX system as a supplement to telehealth musculoskeletal strength assessments.

Methods: In this cross-sectional pilot study, 12 study participants with upper extremity pain and/or weakness underwent strength evaluations for four upper extremity movements using in-person, telehealth, VIRTEPEX, and composite (telehealth plus VIRTEPEX) assessments. The evaluators were blinded to each other's assessments. The primary outcome was feasibility, as determined by participant recruitment, study completion, and safety. The secondary outcome was preliminary evaluation of inter-rater agreement between in-person, telehealth, and VIRTEPEX strength assessments, including κ statistics.

Results: This pilot study had an 80% recruitment rate, a 100% completion rate, and reported no adverse events. In-person and telehealth evaluations achieved highest overall agreement (85.71%), followed by agreements between in-person and composite (75%), in-person and VIRTEPEX (62.5%), and telehealth and VIRTEPEX (62.5%) evaluations. However, for shoulder flexion, agreement between in-person and VIRTEPEX evaluations (78.57%; $\kappa=0.571$, 95% CI 0.183 to 0.960) and in-person and composite evaluations (78.57%; $\kappa=0.571$, 95% CI 0.183 to 0.960) was higher than that between in-person and telehealth evaluations (71.43%; $\kappa=0.429$, 95% CI -0.025 to 0.882).

Conclusions: This study demonstrates the feasibility of asynchronous VIRTEPEX examinations and supports the potential for VIRTEPEX to supplement and add value to standard telehealth platforms. Further studies with an additional development of VIRTEPEX and larger sample sizes for adequate power are warranted.

Keywords: physical examination; telemedicine; tele-health; virtual care; virtual health; telerehabilitation; augmented reality; AR; game; simulation; digital world; virtual environment; motion; strength; force; musculoskeletal; remote examination; exercise; physical examinations; telehealth; cross-sectional; VIRTEPEX; patient; exergame

Introduction

Telemedicine has seen increased utilization in recent years, especially during the COVID-19 pandemic [1]. An important advantage of telemedicine is flexibility, which has led to its incorporation into various medical specialties, including radiology, psychiatry, dermatology, neurology, and cardiology [2]. Telemedicine can be useful for providing health care in difficult-to-reach locations, such as global health settings [3], rural or wilderness areas [4], disaster scenarios [5], and even outer space [6]. Many of these applications use video conferencing technology combined with peripheral examination devices, such as electronic stethoscopes, tele-ophthalmoscopy cameras, video otoscopes, tele-dermatoscopes, digital endoscopes, electronic scales, smartphones, and wearable devices [2]. In the field of rehabilitation medicine, telecommunications technology has enabled telerehabilitation facilitating remote patient interactions [7]. Examples of telerehabilitation applications include replacing in-person visits for chronic lower back pain evaluation and management [8], providing remote synchronous treatment interventions for musculoskeletal conditions [9], and implementing asynchronous rehabilitation programs following total knee replacement surgery [10].

However, existing telemedicine and telerehabilitation platforms have limitations. One major limitation is that physical examinations are difficult to conduct through virtual platforms, as most telemedicine systems are limited to video and audio transmission. Consequently, many physical examination components, including auscultation [11], palpation [2], and manual strength assessments [12] cannot be easily performed through video or audio alone. These assessments may require either another provider at the remote site to assist with physical examination [2] or a peripheral device to collect biometric data. For musculoskeletal examinations, health care providers using existing telemedicine systems may attempt to remotely evaluate patients' range of motion and concentric and eccentric strength by asking them to move a joint or extremity across its full range of motion, raise a limb against gravity, or lift objects of known weight [12]. However, these approaches are limited compared to in-person musculoskeletal examinations and may not always provide reliable information for a comprehensive tele-physical assessment (TelePA) by health care providers.

In recent years, emerging technologies have shown promise in improving reliability and accuracy of TelePA. Wearable biometric sensors can facilitate remote evaluation by transmitting spatiotemporal position, speed, acceleration, gait, force, and haptic data in real time [13-16]. Virtual and augmented reality technologies can immerse patients in simulated environments [13] and facilitate body ownership in virtual settings [17], thereby leading to greater patient

motivation, engagement, and psychosocial benefits during telerehabilitation [18]. Some TelePA systems utilize motion-sensing 3D sensors, such as the Microsoft Kinect system [19], which is a portable, low-cost motion analysis system equipped with RGB+Depth (RGB-D) camera technology [20]. The Kinect system has been studied for various medical applications that involve examining upper and lower extremity function, including assessment of patient movement and function [21], detecting patient walking [22], performing gait analysis [23], assessing balance [24], and mobility monitoring in patients with Parkinson's disease [25].

Methods

Overview

This pilot study aimed to expand the scope of telemedicine and telerehabilitation by assessing the feasibility of the novel Virtual Remote Tele-Physical Examination (VIRTEPEX) system, which combines an augmented reality computer game environment with the Microsoft Kinect system to track motion and estimate force during gameplay [26]. The VIRTEPEX system can be operated asynchronously to evaluate motion and strength and has the potential to augment TelePA during or between telemedicine visits.

VIRTEPEX Game System

VIRTEPEX is a proprietary technology platform that uses the Microsoft Kinect (version 2) RGB-D camera [20] and machine learning software [27]. The technical specifications of VIRTEPEX have been previously published [26]. VIRTEPEX uses noninvasive motion tracking and inverse dynamics to estimate forces for 4 joint movements: shoulder abduction, shoulder flexion, elbow flexion, and wrist extension. As part of the user experience, patients perform the 4 joint movements while playing an augmented reality bowling computer game (Figure 1), which was rendered using Unity LTS (version 2018.4; Unity Technologies). For each joint movement completed, the patient's strength determines the momentum of a virtual bowling ball for each completed joint movement. VIRTEPEX records and transmits force estimates for each joint movement to a health care provider, who can asynchronously assess the patient's strength. During asynchronous evaluation, the health care provider performs the same 4 joint movements through VIRTEPEX; for each joint movement performed by both the patient and provider, VIRTEPEX synthesizes a comparative animation with two virtual bowling balls colliding into one another (Figure 2). In the animation, the momentum of one ball corresponds to the patient's strength for the joint movement, while the other corresponds to the provider's strength. A virtual midline marks the location where the balls would collide if both users applied equal strength; the provider visually estimates the distance between the bowling balls' collision point and the

midline to subjectively estimate the difference in strength between a patient and the provider. In other words, if the balls collide further away from the midline and closer to the

patient, the patient's joint movement strength is judged to be weaker compared to the provider.

Figure 1. The simulated bowling computer game interface [26]. PM&R: Physical Medicine and Rehabilitation. ROM: Range of Motion

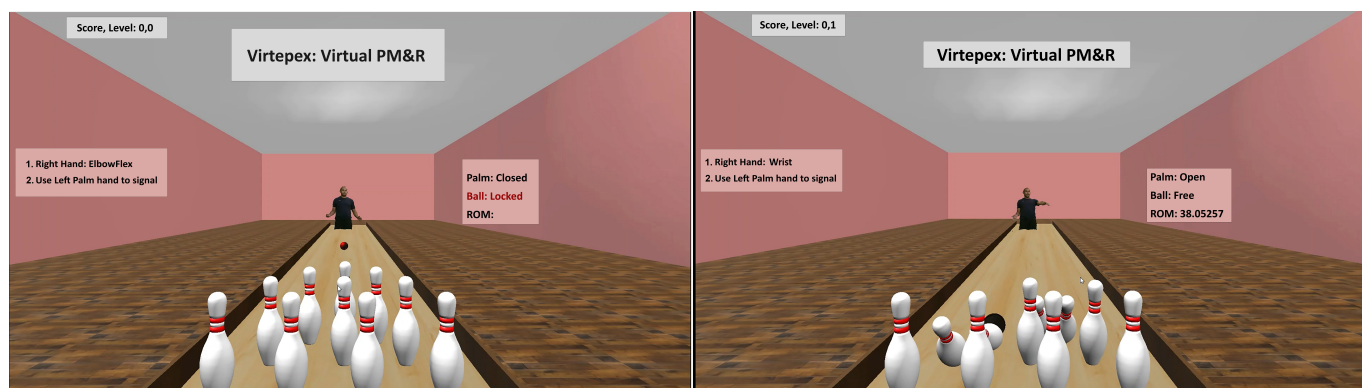
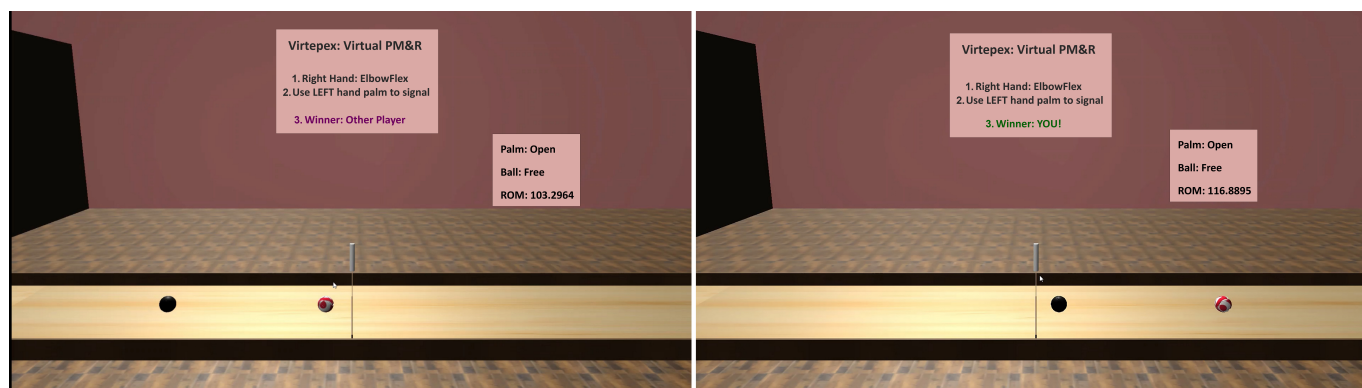


Figure 2. The simulated bowling ball collision animation as viewed by the health care provider for asynchronous remote evaluation of patient's strength [26]. PM&R: Physical Medicine and Rehabilitation. ROM: Range of Motion



Study Sample

This cross-sectional pilot study included patients with impaired upper extremity movements due to any condition causing shoulder, elbow, or wrist pain and/or weakness, with the aim of evaluating movement strength. Adult patients with history of pain and/or weakness in these areas were recruited from a Veterans Affairs hospital's outpatient physical medicine and rehabilitation clinic over one month. Patients scheduled for a therapy (ie, physical, occupational, or kinesiological) for any of the above symptoms were approached for participation; of 15 patients who initially agreed to participate, 3 patients did not show up for their scheduled research visit, leaving 12 patients who were finally recruited. Of these, 10 participants had unilateral upper extremity pain and/or weakness while the other 2 participants had bilateral issues, resulting in a total of 14 evaluations. This number is similar to that previously reported for comparable clinical studies of telerehabilitation technology [7]. Before formal evaluations, participants and evaluators received training on VIRTEPEX and were allowed to interact with VIRTEPEX for some time to acclimate to the system.

Clinical Evaluation

Each participant's strength for 4 upper extremity movements—shoulder abduction, shoulder flexion, elbow flexion, and wrist extension—was assessed by 3 different evaluators blinded to each other's assessments. The first evaluator (a physiatrist with over 20 years of experience) conducted an in-person physical examination; the second evaluator (a medical student adequately trained in physical assessment and strength examination) conducted a synchronous standard telehealth assessment via real-time audio and video of the patient's physical movements; and the third (another trained medical student) conducted an asynchronous VIRTEPEX assessment using the augmented reality bowling computer game alone, without corresponding audio or video of the participants' movements. For every assessment of different joint movements, each evaluator subjectively rated patient strength on a binary scale (normal or impaired) to simplify the evaluation process and subsequent statistical analyses. In addition, the telehealth and VIRTEPEX evaluations were incorporated together into a composite evaluation that aimed to simulate a typical use-case scenario for VIRTEPEX in order to augment a typical telehealth assessment. The composite evaluation included data for wrist extension from telehealth assessments and data for

shoulder abduction, shoulder flexion, and elbow flexion from VIRTEPEX assessments. This approach was formulated by the evaluator using clinical judgement to decide between data from VIRTEPEX and telehealth assessments, a process similar to using clinical judgement to extrapolate information from multiple in-person manual muscle testing data points. During evaluators' training and interactions with VIRTEPEX, it was noted that wrist extension was identified as the most difficult upper extremity movement to capture with the Kinect camera. Therefore, the composite evaluation defaulted to telehealth assessment data for all patients' wrist extension movements.

Study Outcomes

Given this was a pilot study, our primary outcome was to assess for feasibility, based on participant recruitment, study completion, and safety. Recruitment rate was defined as the number of individuals who successfully enrolled and completed the study out of the total patients who were initially approached. Participants completed the study after undergoing each of the 4 upper extremity movement evaluations with each of the 3 evaluators. Safety was measured using the number of adverse events reported during VIRTEPEX evaluation.

The secondary outcome was preliminary assessment of inter-rater agreements, which were calculated for each of the 4 joint movements between (1) in-person and telehealth evaluations, (2) in-person and VIRTEPEX evaluations, (3) in-person and composite evaluations, and (4) telehealth and VIRTEPEX evaluations. Raw percent agreement between two evaluators was calculated as the number of evaluations with agreement divided by the total number of evaluations. κ statistics with 95% CI were calculated using Microsoft Excel for each joint movement to further quantify agreement between in-person and telehealth evaluations, in-person and VIRTEPEX evaluations, in-person and composite evaluations, and telehealth and VIRTEPEX evaluations. κ values were categorized on the following scale used in previous literature [28]: 0.81-1.00 (almost perfect agreement), 0.61-0.80 (excellent agreement), 0.41-0.60 (moderate agreement), 0.21-0.41 (fair agreement), 0.0-0.20 (slight agreement), and <0.00 (no agreement). Overall raw percent agreement across all joint movements was also calculated for each participant.

Ethical Considerations

This study was approved by our institutional research ethics committee (VA Hospital Institutional Review Board, study 1643604). All participants provided written informed consent. Once evaluations were completed, data was deidentified.

Participants in this project were not provided any compensation.

Results

User Statistics

Among 15 patients who were initially approached and agreed to participate in this study, 12 completed the evaluation, while 3 did not attend their scheduled visit. As 2 participants had bilateral upper extremity deficits and were separately evaluated for each side, there were a total of 14 completed evaluations. The study recruitment rate was 80% (N=12), with 100% of the 12 participants completing the study, and no adverse events reported.

For each evaluation method (in-person, telehealth, VIRTEPEX, and composite evaluations), 14 evaluations were conducted. This included assessments of 4 individual joint movements per evaluation, yielding 56 joint assessments per evaluation method and a total of 224 separate joint assessments across all methods. Among in-person evaluations, 38 joint assessments were rated as normal and 18 as impaired. Telehealth evaluations rated 40 joint assessments as normal and 16 as impaired. VIRTEPEX evaluations identified 37 joint assessments as normal and 19 as impaired, while composite evaluations identified 44 joint assessments as normal and 12 as impaired.

Evaluation Outcomes

In-Person Versus Telehealth

Overall raw agreements for individual joint movements ranged from 71.43% to 100%, with 85.71% for shoulder abduction, 71.43% for shoulder flexion, 100% for elbow flexion, and 85.71% for wrist extension during comparison between in-person and telehealth evaluations (Table 1). κ values exhibited moderate to almost perfect agreement, ranging from $\kappa=0.429$ to $\kappa=1.000$. Substantial agreement ($\kappa=0.720$, 95% CI 0.375 to 1.000) was observed for shoulder abduction, moderate agreement ($\kappa=0.429$, 95% CI -0.025 to 0.882) for shoulder flexion, almost perfect agreement ($\kappa=1.000$, 95% CI 1.000 to 1.000) for elbow flexion, and moderate agreement ($\kappa=0.440$, 95% CI -0.155 to 1.000) for wrist extension. Based on κ values, the highest agreement was observed for elbow flexion (almost perfect with 100.00% raw agreement; $\kappa=1.000$, 95% CI 1.000 to 1.000), and the lowest for shoulder flexion (moderate agreement with 71.43% raw agreement, $\kappa=0.429$, 95% CI -0.025 to 0.882). Overall agreement between in-person and telehealth assessments was 85.71%, varying from 50% to 100% per participant.

Table 1. Raw percent agreements, κ statistics, and CI for in-person versus telehealth evaluations.

Variables	Joints and movements				Raw agreement (%)
	Shoulder abduction	Shoulder flexion	Elbow flexion	Wrist extension	
Research evaluation ^a					
1	A ^b	A	A	A	100

Variables	Joints and movements				Raw agreement (%)
	Shoulder abduction	Shoulder flexion	Elbow flexion	Wrist extension	
2	D ^c	A	A	A	75
3	A	A	A	A	100
4	A	A	A	A	100
5	D	D	A	A	50
6	A	D	A	D	50
7	A	A	A	A	100
8	A	D	A	A	75
9	A	A	A	A	100
10	A	A	A	D	75
11	A	A	A	A	100
12	A	A	A	A	100
13	A	A	A	A	100
14	A	D	A	A	75
Overall agreement	85.71	71.43	100	85.71	85.71
κ statistic (95% CI)	0.720 (0.375 to 1.000)	0.429 (-0.025 to 0.882)	1.000 (1.000 to 1.000)	0.440 (-0.155 to 1.000)	— ^d

^aIndividual participants' responses.

^bA: Agree.

^cD: Disagree.

^dNot applicable.

In-Person Versus VIRTEPEX

Overall raw agreements for individual joint movements ranged from 35.71% to 78.57%, with 57.14% for shoulder abduction, 78.57% for shoulder flexion, 78.57% for elbow flexion, and 35.71% for wrist extension during comparison between in-person evaluations and VIRTEPEX evaluations (Table 2). κ values exhibited slight to moderate agreement, ranging from $\kappa=0.060$ to $\kappa=0.571$. Fair agreement was observed ($\kappa=0.222$, 95% CI -0.073 to 0.518) for shoulder abduction, moderate agreement ($\kappa=0.571$, 95% CI 0.183 to

0.960) for shoulder flexion, fair agreement ($\kappa=0.276$, 95% CI -0.332 to 0.883) for elbow flexion, and slight agreement ($\kappa=0.060$, 95% CI -0.065 to 0.185) for wrist extension. Based on κ values, the highest agreement was observed for shoulder flexion (moderate agreement with 78.57% raw agreement; $\kappa=0.571$, 95% CI 0.183 to 0.960), and lowest for wrist extension (slight agreement with 35.71% raw agreement; $\kappa=0.060$, 95% CI -0.065 to 0.185). Overall agreement between in-person evaluations and VIRTEPEX was 62.5% , varying from 50% to 75% per participant.

Table 2. Raw percent agreements, κ statistics, and confidence intervals (CI) for in-person versus VIRTEPEX evaluations.

Variables	Joints and movements				Raw agreement (%)
	Shoulder abduction	Shoulder flexion	Elbow flexion	Wrist extension	
Research evaluation ^a					
1	D ^b	A ^c	A	A	75
2	D	A	A	D	50
3	A	A	A	D	75
4	A	D	A	A	75
5	D	D	A	A	50
6	D	A	D	A	50
7	A	D	A	D	50
8	D	A	A	A	75
9	A	A	D	D	50
10	A	A	A	D	75
11	A	A	A	D	75
12	A	A	A	D	75
13	A	A	D	D	50
14	D	A	A	D	50

Variables	Joints and movements				Raw agreement (%)
	Shoulder abduction	Shoulder flexion	Elbow flexion	Wrist extension	
Overall agreement (%)	57.14	78.57	78.57	35.71	62.50
κ statistic (95% CI)	0.222 (-0.073 to 0.518)	0.571 (0.183 to 0.960)	0.276 (-0.332 to 0.883)	0.060 (-0.065 to 0.185)	— ^d

^aIndividual participants' responses.

^bD: Disagree.

^cA: Agree.

^dNot applicable.

In-Person Versus Composite Evaluations

Overall raw agreements for individual joint movements ranged from 57.14% to 85.71%, with 57.14% for shoulder abduction, 78.57% for shoulder flexion, 78.57% for elbow flexion, and 85.71% for wrist extension during comparison between in-person and composite evaluations (Table 3). The κ values exhibited fair to moderate agreement, ranging from $\kappa=0.222$ to $\kappa=0.571$. Fair agreement ($\kappa=0.222$, 95% CI -0.073 to 0.518) was observed for shoulder abduction, moderate agreement ($\kappa=0.571$, 95% CI 0.183 to 0.960) for

shoulder flexion, fair agreement ($\kappa=0.276$, 95% CI -0.332 to 0.883) for elbow flexion, and moderate agreement ($\kappa=0.440$, 95% CI -0.155 to 1.000) for wrist extension. Based on κ values, the highest agreement was observed for shoulder flexion (moderate agreement with 78.57% raw agreement; $\kappa=0.571$, 95% CI 0.183 to 0.960), and lowest for shoulder abduction (fair agreement with 57.14% raw agreement; $\kappa=0.222$, 95% CI -0.073 to 0.518). Overall agreement between in-person and composite evaluations was 75%, varying from 25% to 100% per participant.

Table 3. Raw percent agreements, κ statistics, and CI for in-person versus composite evaluations.

Variables	Joints and movements				Raw agreement (%)
	Shoulder abduction	Shoulder flexion	Elbow flexion	Wrist extension	
Research evaluation ^a					
1	D ^b	A ^c	A	A	75
2	D	A	A	A	75
3	A	A	A	A	100
4	A	D	A	A	75
5	D	D	A	A	50
6	D	A	D	D	25
7	A	D	A	A	75
8	D	A	A	A	75
9	A	A	D	A	75
10	A	A	A	D	75
11	A	A	A	A	100
12	A	A	A	A	100
13	A	A	D	A	75
14	D	A	A	A	75
Overall agreement (%)	57.14	78.57	78.57	85.71	75
κ statistic (95% CI)	0.222 (-0.073 to 0.518)	0.571 (0.183 to 0.960)	0.276 (-0.332 to 0.883)	0.440 (-0.155 to 1.000)	— ^d

^aIndividual participants' responses.

^bD: Disagree.

^cA: Agree.

^dNot applicable.

Telehealth Versus VIRTEPEX

Overall raw agreements for individual joint movements ranged from 35.71% to 78.57%, with 71.43% for shoulder abduction, 64.29% for shoulder flexion, 78.57% for elbow flexion, and 35.71% for wrist extension during the comparison between telehealth and VIRTEPEX (Table 4). κ values exhibited poor to fair agreement, ranging from $\kappa=-0.033$ to $\kappa=0.364$. Fair agreement ($\kappa=0.364$, 95% CI -0.043 to 0.770) was observed for shoulder abduction; slight

agreement ($\kappa=0.186$, 95% CI -0.342 to 0.714) for shoulder flexion, fair agreement ($\kappa=0.276$, 95% CI -0.332 to 0.883) for elbow flexion, and poor agreement ($\kappa=-0.033$, 95% CI -0.356 to 0.291) for wrist extension. Based on κ values, the highest agreement was observed for shoulder abduction (fair agreement with 71.43% raw agreement; $\kappa=0.364$, 95% CI -0.043 to 0.770), and lowest for wrist extension (poor agreement with 35.71% raw agreement; $\kappa=-0.033$, 95% CI -0.356 to 0.291). Overall agreement was 62.5%, varying from 0% to 100% per participant.

Table 4. Raw percent agreements, κ statistics, and CI for telehealth versus VIRTEPEX evaluations.

Variables	Joints and movements				Raw agreement (%)
	Shoulder abduction	Shoulder flexion	Elbow flexion	Wrist extension	
Research evaluation ^a					
1	D ^b	A ^c	A	A	75
2	A	A	A	D	75
3	A	A	A	D	75
4	A	D	A	A	75
5	A	A	A	A	100
6	D	D	D	D	0
7	A	D	A	D	50
8	D	D	A	A	50
9	A	A	D	D	50
10	A	A	A	A	100
11	A	A	A	D	75
12	A	A	A	D	75
13	A	A	D	D	50
14	D	D	A	D	25
Overall agreement (%)	71.43	64.29	78.57	35.71	62.5
κ statistic (95% CI)	0.364 (−0.043 to 0.770)	0.186 (−0.342 to 0.714)	0.276 (−0.332 to 0.883)	−0.033 (−0.356 to 0.291)	− ^d

^aIndividual participants' responses.

^bD: Disagree.

^cA: Agree.

^dNot applicable.

Discussion

Principal Results

This pilot study suggests that using VIRTEPEX to supplement telehealth strength assessment is feasible and safe. Further, VIRTEPEX demonstrates sufficiently acceptable levels of inter-rater agreement with in-person examination to warrant further evaluation for clinical use.

The inclusion of the composite assessment in this study illustrates how VIRTEPEX could be integrated into an existing clinical workflow. VIRTEPEX enables asynchronous collection of strength assessment data, which can help inform subsequent synchronous telehealth examinations. These synchronous examinations could be more efficient by focusing on movements identified as deficient via VIRTEPEX. In clinical practice, the providers could use clinical judgment to create a comprehensive assessment based on datapoints from different examination platforms. The composite evaluation in this study aims to represent this process and evaluate the utility of VIRTEPEX, and findings suggest that VIRTEPEX has the potential to augment remote strength assessments.

Among telehealth, VIRTEPEX, and composite evaluations, telehealth evaluation showed the highest overall raw agreement with in-person evaluation (85.71%). This was higher than raw percent agreements between in-person and VIRTEPEX evaluations (62.5%), in-person and composite

evaluations (75%), and telehealth and VIRTEPEX evaluations (62.5%). Previous studies on inter-rater reliability for upper extremity musculoskeletal examinations have shown a wide but comparable range of agreement values. For instance, one study demonstrated raw inter-rater agreement values ranging from 66.67% to 98.9% for in-person upper extremity examinations performed by different evaluators [28], while another study comparing telemedicine-based shoulder examinations with in-person evaluations exhibited raw agreement values ranging from 46.7% to 83.7% [29].

Statistically significant agreement was observed between in-person and telehealth evaluations for shoulder abduction and elbow flexion (substantial agreement: $\kappa=0.720$, 95% CI 0.375 to 1.000 for shoulder abduction; almost perfect agreement: $\kappa=1.000$, 95% CI 1.000 to 1.000 for elbow flexion). Statistically significant, moderate agreement was observed between in-person and VIRTEPEX ($\kappa=0.571$, 95% CI 0.183 to 0.960) and in-person and composite evaluations ($\kappa=0.571$, 95% CI 0.183 to 0.960) for shoulder flexion. However, the agreement between in-person and telehealth evaluations for shoulder flexion was not statistically significant ($\kappa=0.429$, 95% CI −0.025 to 0.882). In contrast, the statistical significance of the agreements for shoulder flexion between in-person and VIRTEPEX evaluations, as well as between in-person and composite evaluations was particularly notable. These results suggest that while telehealth may have greater overall agreement with in-person evaluation and be better-suited for evaluating shoulder abduction and elbow flexion strength, VIRTEPEX

and composite evaluations may be superior for assessing shoulder flexion strength in remote settings.

Interestingly, none of the remote evaluation methods demonstrated statistically significant agreement with in-person evaluations for wrist extension strength assessment. This suggests that further modifications to both existing telehealth technology and VIRTEPEX are needed to improve remote evaluation of patients with wrist weakness and/or pain. Additionally, while telehealth and VIRTEPEX evaluations showed an overall agreement of 62.5%, the lack of statistically significant κ values for the 4 joint movements suggests that the observed agreement may have been due to chance. Alternatively, it may indicate that telehealth and VIRTEPEX are better suited for assessing joint movements differently, resulting in distinct data patterns across these evaluation platforms.

Limitations

This pilot study had several limitations. First, given that using VIRTEPEX requires reliable internet connection, this system would have limited utility for patients without internet access. Second, the sample size was small (14 evaluations) and all participants were veterans, reducing generalizability of findings to nonveteran populations. As this was a pilot study focused on feasibility, the small sample size provided less than 80% power—corresponding to approximately 22-30 evaluations—to distinguish moderate inter-rater agreement ($\kappa=0.6$) in a two-sided test [30]. A larger, fully powered study would be necessary to better quantify agreement across different evaluation methods. Additionally, VIRTEPEX is currently still under development and is not yet available for clinical practice, as further work is needed to assess and improve the system's ease of use, cost-effectiveness, clinical outcomes, and user satisfaction. Similarly, the subjective binary strength grading used in this study was applied for simplicity and convenience during the evaluation process and statistical analyses; however, it is not applied as a standard in clinical practice.

Acknowledgments

Research from this study was previously presented as poster presentations and published as abstracts for the following conferences: (1) the 2022 Association for Computing Machinery Designing Interactive Systems Conference 26, (2) 2022 American Academy of Physical Medicine & Rehabilitation Annual Assembly 31, (3) 2023 Association of Academic Physiatrists Annual Meeting 32

Disclaimer

VIRTEPEX is a proprietary device/product that has not yet been approved for use by the FDA for any purpose, is not labeled for use, and is still under investigation.

Data Availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Authors' Contributions

Conceptualization: RW, KC, SB, NK, KD, BP, TA

Data curation: RW, KC, TA

Development of VIRTEPEX: NK, KD, BP, TA

Formal analysis: RW

Methodology: RW, KC, SB, NK, KD, BP, TA

Future Directions

In this pilot study, VIRTEPEX was only used to remotely assess upper extremity movement strength. However, future development could enable VIRTEPEX to remotely assess strength for joint movements in other areas of the body, such as hips, knees, ankles, torso, and neck, facilitating a more comprehensive TelePA. Additional development could potentially enable VIRTEPEX to evaluate multi-joint movements, gait, and/or balance. Additionally, given that the VIRTEPEX system did not evaluate pain experienced by patients during movements, another potential feature to implement would be pain assessment.

Other potential steps could help improve appeal and adoption of VIRTEPEX. Although VIRTEPEX is currently optimized for use with Microsoft Kinect, adapting it to other devices could facilitate wider adoption. Further, streamlining the user interface and incorporating more game design elements could simplify the user experience and improve user engagement during telerehabilitation. Finally, VIRTEPEX could be expanded beyond traditional telehealth settings to other clinical applications, such as remote physiological monitoring for patients needing recurrent physical assessments.

Conclusions

This study supports the feasibility of VIRTEPEX as a supplement to telehealth and demonstrates that VIRTEPEX can achieve moderate agreement with in-person evaluations. Notably, VIRTEPEX had greater agreement with in-person evaluations than telehealth for shoulder flexion, suggesting its potential to enhance existing telehealth technologies. Further technological developments to VIRTEPEX, combined with more adequately powered studies, can better evaluate the effectiveness and accuracy of VIRTEPEX-supplemented TelePA.

Supervision: TA

Writing – original draft: RW, KC, SB

Writing – review & editing: TA

Conflicts of Interest

None declared.

References

1. Tenforde AS, Iaccarino MA, Borgstrom H, et al. Telemedicine during COVID-19 for outpatient sports and musculoskeletal medicine physicians. *PM R*. Sep 2020;12(9):926-932. [doi: [10.1002/pmrj.12422](https://doi.org/10.1002/pmrj.12422)] [Medline: [32424977](https://pubmed.ncbi.nlm.nih.gov/32424977/)]
2. Weinstein RS, Krupinski EA, Doarn CR. Clinical examination component of telemedicine, telehealth, mHealth, and connected health medical practices. *Med Clin North Am*. May 2018;102(3):533-544. [doi: [10.1016/j.mcna.2018.01.002](https://doi.org/10.1016/j.mcna.2018.01.002)] [Medline: [29650074](https://pubmed.ncbi.nlm.nih.gov/29650074/)]
3. Dietrich D, Dekova R, Davy S, Fahrni G, Geissbühler A. Applications of space technologies to global health: scoping review. *J Med Internet Res*. Jun 27, 2018;20(6):e230. [doi: [10.2196/jmir.9458](https://doi.org/10.2196/jmir.9458)] [Medline: [29950289](https://pubmed.ncbi.nlm.nih.gov/29950289/)]
4. Ting L, Wilkes M. Telemedicine for patient management on expeditions in remote and austere environments: a systematic review. *Wilderness Environ Med*. Mar 2021;32(1):102-111. [doi: [10.1016/j.wem.2020.09.009](https://doi.org/10.1016/j.wem.2020.09.009)] [Medline: [33423896](https://pubmed.ncbi.nlm.nih.gov/33423896/)]
5. Merrell RC, Cone SW, Rafiq A. Telemedicine in extreme conditions: disasters, war, remote sites. *Stud Health Technol Inform*. 2008;131:99-116. [Medline: [18305326](https://pubmed.ncbi.nlm.nih.gov/18305326/)]
6. Grigoriev AI, Orlov OI. Telemedicine and spaceflight. *Aviat Space Environ Med*. Jul 2002;73(7):688-693. [Medline: [12137107](https://pubmed.ncbi.nlm.nih.gov/12137107/)]
7. Schutte J, Gales S, Filippone A, Saptono A, Parmanto B, McCue M. Evaluation of a telerehabilitation system for community-based rehabilitation. *Int J Telerehabil*. 2012;4(1):15-24. [doi: [10.5195/ijt.2012.6092](https://doi.org/10.5195/ijt.2012.6092)] [Medline: [25945193](https://pubmed.ncbi.nlm.nih.gov/25945193/)]
8. Jansen-Kosterink S, Huis in 't Veld R, Wever D, Hermens H, Vollenbroek-Hutten M. Introducing remote physical rehabilitation for patients with chronic disorders by means of telemedicine. *Health Technol (Berl)*. Jul 1, 2015;5(2):83-90. [doi: [10.1007/s12553-015-0111-5](https://doi.org/10.1007/s12553-015-0111-5)]
9. Cottrell MA, Galea OA, O'Leary SP, Hill AJ, Russell TG. Real-time telerehabilitation for the treatment of musculoskeletal conditions is effective and comparable to standard practice: a systematic review and meta-analysis. *Clin Rehabil*. May 2017;31(5):625-638. [doi: [10.1177/0269215516645148](https://doi.org/10.1177/0269215516645148)] [Medline: [27141087](https://pubmed.ncbi.nlm.nih.gov/27141087/)]
10. Bini SA, Mahajan J. Clinical outcomes of remote asynchronous telerehabilitation are equivalent to traditional therapy following total knee arthroplasty: a randomized control study. *J Telemed Telecare*. Feb 2017;23(2):239-247. [doi: [10.1177/1357633X16634518](https://doi.org/10.1177/1357633X16634518)] [Medline: [26940798](https://pubmed.ncbi.nlm.nih.gov/26940798/)]
11. Kang SH, Joe B, Yoon Y, Cho GY, Shin I, Suh JW. Cardiac auscultation using smartphones: pilot study. *JMIR Mhealth Uhealth*. Feb 28, 2018;6(2):e49. [doi: [10.2196/mhealth.8946](https://doi.org/10.2196/mhealth.8946)] [Medline: [29490899](https://pubmed.ncbi.nlm.nih.gov/29490899/)]
12. Tanaka MJ, Oh LS, Martin SD, Berkson EM. Telemedicine in the era of COVID-19: the virtual orthopaedic examination. *J Bone Joint Surg Am*. Jun 17, 2020;102(12). [doi: [10.2106/JBJS.20.00609](https://doi.org/10.2106/JBJS.20.00609)] [Medline: [32341311](https://pubmed.ncbi.nlm.nih.gov/32341311/)]
13. Borresen A, Wolfe C, Lin CK, et al. Usability of an immersive augmented reality based telerehabilitation system with haptics (ARTESH) for synchronous remote musculoskeletal examination. *Int J Telerehabil*. 2019;11(1):23-32. [doi: [10.5195/ijt.2019.6275](https://doi.org/10.5195/ijt.2019.6275)] [Medline: [31341544](https://pubmed.ncbi.nlm.nih.gov/31341544/)]
14. Rodríguez-Martín D, Pérez-López C, Samà A, et al. A waist-worn inertial measurement unit for long-term monitoring of Parkinson's disease patients. *Sensors (Basel)*. Apr 11, 2017;17(4):827. [doi: [10.3390/s17040827](https://doi.org/10.3390/s17040827)] [Medline: [28398265](https://pubmed.ncbi.nlm.nih.gov/28398265/)]
15. Zhang S, Guo S, Gao B, Hirata H, Ishihara H. Design of a novel telerehabilitation system with a force-sensing mechanism. *Sensors (Basel)*. May 19, 2015;15(5):11511-11527. [doi: [10.3390/s150511511](https://doi.org/10.3390/s150511511)] [Medline: [25996511](https://pubmed.ncbi.nlm.nih.gov/25996511/)]
16. Piqueras M, Marco E, Coll M, et al. Effectiveness of an interactive virtual telerehabilitation system in patients after total knee arthroplasty: a randomized controlled trial. *J Rehabil Med*. Apr 2013;45(4):392-396. [doi: [10.2340/16501977-1119](https://doi.org/10.2340/16501977-1119)] [Medline: [23474735](https://pubmed.ncbi.nlm.nih.gov/23474735/)]
17. Matamala-Gomez M, Bottiroli S, Realdon O, et al. Telemedicine and virtual reality at time of COVID-19 pandemic: an overview for future perspectives in neurorehabilitation. *Front Neurol*. 2021;12:646902. [doi: [10.3389/fneur.2021.646902](https://doi.org/10.3389/fneur.2021.646902)] [Medline: [33841313](https://pubmed.ncbi.nlm.nih.gov/33841313/)]
18. Howard MC. A meta-analysis and systematic literature review of virtual reality rehabilitation programs. *Comput Human Behav*. May 1, 2017;70:317-327. [doi: [10.1016/j.chb.2017.01.013](https://doi.org/10.1016/j.chb.2017.01.013)]
19. Timmi A, Coates G, Fortin K, et al. Accuracy of a novel marker tracking approach based on the low-cost Microsoft Kinect v2 sensor. *Med Eng Phys*. Sep 2018;59:63-69. [doi: [10.1016/j.medengphy.2018.04.020](https://doi.org/10.1016/j.medengphy.2018.04.020)] [Medline: [29983277](https://pubmed.ncbi.nlm.nih.gov/29983277/)]

20. Gaddam SPR, Chippa MK, Sastry S, Ange A, Berki V, Davis BL. Estimating forces during exercise activity using non-invasive Kinect camera. Presented at: 2015 International Conference on Computational Science and Computational Intelligence (CSCI); Dec 7-9, 2015; Las Vegas, NV, USA. [doi: [10.1109/CSCI.2015.126](https://doi.org/10.1109/CSCI.2015.126)]
21. Capecchi M, Ceravolo MG, Ferracuti F, et al. Accuracy evaluation of the Kinect v2 sensor during dynamic movements in a rehabilitation scenario. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Aug 16-20 2016;2016:5409-5412. [doi: [10.1109/EMBC.2016.7591950](https://doi.org/10.1109/EMBC.2016.7591950)] [Medline: [28269481](https://pubmed.ncbi.nlm.nih.gov/28269481/)]
22. Liu L, Mehrotra S. Patient walk detection in hospital room using Microsoft Kinect v2. Presented at: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); Aug 16-20, 2016:4395-4398; Orlando, FL, USA. [doi: [10.1109/EMBC.2016.7591701](https://doi.org/10.1109/EMBC.2016.7591701)]
23. Latorre J, Colomer C, Alcañiz M, Llorens R. Gait analysis with the Kinect v2: normative study with healthy individuals and comprehensive study of its sensitivity, validity, and reliability in individuals with stroke. J Neuroeng Rehabil. Jul 26, 2019;16(1):97. [doi: [10.1186/s12984-019-0568-y](https://doi.org/10.1186/s12984-019-0568-y)] [Medline: [31349868](https://pubmed.ncbi.nlm.nih.gov/31349868/)]
24. Ayed I, Jaume-i-Capó A, Martínez-Bueso P, Mir A, Moyà-Alcover G. Balance measurement using Microsoft Kinect v2: towards remote evaluation of patient with the functional reach test. Appl Sci (Basel). 2021;11(13):6073. [doi: [10.3390/app11136073](https://doi.org/10.3390/app11136073)] [Medline: [6073](https://pubmed.ncbi.nlm.nih.gov/6073/)]
25. Amini A, Banitsas K, Young WR. Kinect4FOG: monitoring and improving mobility in people with Parkinson's using a novel system incorporating the Microsoft Kinect v2. Disabil Rehabil Assist Technol. Aug 2019;14(6):566-573. [doi: [10.1080/17483107.2018.1467975](https://doi.org/10.1080/17483107.2018.1467975)] [Medline: [29790385](https://pubmed.ncbi.nlm.nih.gov/29790385/)]
26. Khargonkar N, Desai K, Prabhakaran B, Annaswamy T. Virteplex: virtual remote tele-physical examination system. Presented at: Designing Interactive Systems Conference Proceedings; Jun 13-17, 2022; Virtual Event Australia. [doi: [10.1145/3532106.3533486](https://doi.org/10.1145/3532106.3533486)]
27. Shotton J, Fitzgibbon A, Cook M, et al. Real-time human pose recognition in parts from single depth images. Presented at: Conference on Computer Vision and Pattern Recognition Proceedings. 20-25 June, 2011.[doi: [10.1109/cvpr.2011.5995316](https://doi.org/10.1109/cvpr.2011.5995316)]
28. Smith CK, Bonauto DK, Silverstein BA, Wilcox D. Inter-rater reliability of physical examinations in a prospective study of upper extremity musculoskeletal disorders. J Occup Environ Med. Oct 2010;52(10):1014-1018. [doi: [10.1097/JOM.0b013e3181f4396b](https://doi.org/10.1097/JOM.0b013e3181f4396b)] [Medline: [20881623](https://pubmed.ncbi.nlm.nih.gov/20881623/)]
29. Wang G, Fiedler AK, Warth RJ, Bailey L, Shupe PG, Gregory JM. Reliability and accuracy of telemedicine-based shoulder examinations. J Shoulder Elbow Surg. Aug 1, 2022;31(8):e369-e375. [doi: [10.1016/j.jse.2022.04.005](https://doi.org/10.1016/j.jse.2022.04.005)]
30. Sim J, Wright CC. The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Phys Ther. Mar 2005;85(3):257-268. [Medline: [15733050](https://pubmed.ncbi.nlm.nih.gov/15733050/)]
31. Wu R, Chakka K, Belko S, 2022 American Academy of Physical Medicine & Rehabilitation Annual Assembly. Virtual remote tele-physical examination (VIRTEPEX) system user validation. PM R. Oct 2022;14(S1). [doi: [10.1002/pmrj.12913](https://doi.org/10.1002/pmrj.12913)]
32. Chakka K, Wu R, Belko S. User validation of the virtual remote tele-physical examination (VIRTEPEX) system. Am J Phys Med Rehabil. 2023;102(4S):A1-A234. [doi: [10.1097/PHM.0000000000002184](https://doi.org/10.1097/PHM.0000000000002184)]

Abbreviations

RGB-D: RGB+Depth

TelePA: tele-physical assessment

VA: Veterans Affairs

VIRTEPEX: Virtual Remote Tele-Physical Examination

Edited by Maurizio Schmid; peer-reviewed by Lucia Mosca, Thomas Davidson; submitted 17.02.2024; final revised version received 09.12.2024; accepted 22.12.2024; published 05.02.2025

Please cite as:

Wu R, Chakka K, Belko S, Khargonkar N, Desai K, Prabhakaran B, Annaswamy T

Comparing In-Person, Standard Telehealth, and Remote Musculoskeletal Examination With a Novel Augmented Reality Exercise Game System: Pilot Cross-Sectional Comparison Study

JMIR Serious Games 2025;13:e57443

URL: <https://games.jmir.org/2025/1/e57443>

doi: [10.2196/57443](https://doi.org/10.2196/57443)

© Richard Wu, Keerthana Chakka, Sara Belko, Ninad Khargonkar, Kevin Desai, Balakrishnan Prabhakaran, Thiru Annaswamy. Originally published in JMIR Serious Games (<https://games.jmir.org>), 05.02.2025. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Serious Games, is properly cited. The complete bibliographic information, a link to the original publication on <https://games.jmir.org>, as well as this copyright and license information must be included.